CLAIMS

1. A video image object recognizing apparatus comprising:

input means for inputting video image data and image capturing information which is information for determining an area where an image will be captured;

5   storage means for storing positional information which is information representing the position of an object and visual feature information which is information representing a numerical value of a visual feature of the object, that are connected to each other; and

object recognizing means for recognizing an object contained in 10   a video image based on the input video image data;

wherein said object recognizing means comprises:

estimating means for estimating an area where an image will be captured based on the image capturing information;

matching means for matching the area where an image will be 15   captured to a position represented by the positional information of the object stored in said storage means;

partial video image extracting means for extracting partial video image data which is either video image data of a partial area of the video image based on the video image data or is video image data of the entire video im-20   age, from the input video image;

visual feature information setting means for generating visual feature information of the partial video image data;

similarity calculating means for comparing the visual feature information of the partial video image data and the visual feature information of

25   the object stored in said storage means with each other to calculate a similarity

therebetween; and

decision means for determining whether or not an object is pre-

sent in the video image, based on the input video image data, which is based

on the result of matching by said matching means and on the result of the cal-

30   culated similarity.

2. A video image annotation applying apparatus comprising:

input means for inputting video image data and image capturing

information which is information for determining an area where an image will be

captured;

5   storage means for storing positional information which is infor-

mation representing the position of an object, visual feature information which

is information representing a numerical value of a visual feature of the object,

and additional information which is information added to the object, that are

connected to each other; and

10   object recognizing means for associating an object contained in

a video image based on the input video image data with the additional informa-

tion;

wherein said object recognizing means comprises:

estimating means for estimating an area where an image will be

15   captured based on the image capturing information;

matching means for matching the area where an image will be

captured to a position represented by the positional information of the object

stored in said storage means;

partial video image extracting means for extracting partial video

20   image data which is either video image data of a partial area of the video image

based on the video image data or is video image data of the entire video image, from the input video image;

visual feature information setting means for generating visual feature information of the partial video image data;

25   similarity calculating means for comparing the visual feature information of the partial video image data and the visual feature information of the object stored in said storage means with each other to calculate a similarity therebetween; and

decision means for identifying an object which is contained in

30   the video image based on the input video image data, and which is based on the result of the matching by said matching means and the calculated similarity, and for associating the identified object and the additional information stored in said storage means with each other.

3. The video image annotation applying apparatus according to claim 2, wherein said object recognizing means includes:

presence probability calculating means for calculating an presence probability which is the probability that an object is contained in the video

5   image, based on the area where an image will be captured and the positional information of the object stored in the storage means; and

wherein said decision means identifies an object which is contained in the video image based on the calculated presence probability and similarity, and associates the identified object and the additional information

10   stored in said storage means with each other.

4. The video image annotation applying apparatus according to claim 3, wherein said partial video image extracting means identifies a range within which the object is positioned in the video image based on the positional

information of the object stored in the storage means, and extracts partial video

5   image data from the identified range.

        5. The video image annotation applying apparatus according to claim 2, wherein said object recognizing means includes:

        candidate object searching means for extracting a candidate object, which is an object present in the area where an image will be captured,

5   based on the area where an image will be captured and the positional information; and

        wherein said similarity calculating means compares the visual feature information of the partial video image data and the visual feature information of a candidate object stored in said storage means with each other to

10   calculate a similarity therebetween.

        6. The video image annotation applying apparatus according to claim 5, wherein said partial video image extracting means identifies a range within which the object is positioned in the video image based on the positional information of the candidate object stored in the storage means, and extracts

5   partial video image data from the identified range.

        7. The video image annotation applying apparatus according to claim 2, further comprising:

        display means for displaying a video image; and

        display position determining means for indicating a position to

5   display the additional information associated with the object contained in the video image and for displaying the additional information that is superimposed on the video image.

        8. The video image annotation applying apparatus according to claim 2, further comprising:

annotation result storage means for storing the additional information and the object contained in the video image in association with each

5   other.

9. The video image annotation applying apparatus according to claim 2, wherein said partial video image extracting means has a function to arbitrarily change the shape and size of the area of a video image based on the extracted partial video image data.

10. The video image annotation applying apparatus according to claim 2, wherein said partial video image extracting means extracts partial video image data in the area of a video image which matches one or a combination of conditions including luminance information, color information, shape

5   information, texture information, and size information.

11. The video image annotation applying apparatus according to claim 10, wherein if said partial video image extracting means extracts partial video image data from a video image which matches a combination of each condition, then said partial video image extracting means determines an impor-

5   tance of said condition and extracts partial video image data based on the result of the matching by said matching means and the visual feature information of the object stored in the storage means.

12. The video image annotation applying apparatus according to claim 2, wherein the visual feature information of the object stored in the storage means comprises a template video image which is a video image having a visual feature similar to the object.

13. The video image annotation applying apparatus according to claim 2, wherein the visual feature information of the object stored in the storage means comprises one or more items of color information, shape informa-

tion, texture information, and size information, and the visual feature informa-

5   tion of the partial video image data generated by said visual feature information

setting means comprises one or more items of color information, shape infor-

mation, texture information, and size information.

14. The video image annotation applying apparatus according to

claim 2, wherein the positional information of the object stored in said storage

means comprises information for identifying the position of one of the vertexes,

a central point, or a center of gravity of a three-dimensional shape which ap-

5   proximates a three-dimensional shape of solid geometry including a cone, a

cylinder, a cube, or the like which is similar to the object.

15. The video image annotation applying apparatus according to

claim 2, wherein the positional information of the object stored in said storage

means comprises information for identifying the position of at least one of the

vertexes of a three-dimensional shape which approximates the object having

5   polygonal surfaces.

16. The video image annotation applying apparatus according to

claim 2, wherein the positional information of the object stored in said storage

means comprises information for identifying the position of a vertex which is

highest of all the vertexes of the object.

17. The video image annotation applying apparatus according to

claim 2, wherein the positional information of the object stored in said storage

means comprises information for identifying the position of the object according

to a latitude, a longitude, and an altitude.

18. The video image annotation applying apparatus according to

claim 2, wherein said storage means stores in a hierarchical pattern common

additional information based on a concept common to additional information

associated respectively with a plurality of objects or stores common additional

5   information based on a concept common to a plurality of items of common additional information, and said decision means determines whether there is common additional information corresponding to additional information or common additional information of an object whose image is captured, and, if there is such common additional information, associates the object with the common

10   additional information.

19. The video image annotation applying apparatus according to claim 2, wherein said image capturing information includes captured date and time information which is information for identifying a captured date and time, said storage means stores visual feature information depending on the cap-

5   tured date and time, and said similarity calculating means compares the visual feature information of the partial video image data and the visual feature information depending on the captured date and time identified by the captured date and time information with each other to calculate a similarity therebetween.

20. The video image annotation applying apparatus according to claim 10, wherein said partial video image extracting means divides areas from said input video image data and extracts the divided areas as said partial video image data.

21. The video image annotation applying apparatus according to claim 20, wherein said partial video image extracting means combines the divided areas into said partial video image data.

22. The video image annotation applying apparatus according to claim 21, wherein said partial video image extracting means generates the par-

tial video image data by hierarchically evaluating a combination of said divided areas.

23. The video image annotation applying apparatus according to claim 22, wherein said partial video image extracting means uses only a number of areas whose similarity is high for subsequent combination from the combination of areas in hierarchically evaluating the combination of said divided ar-

5 eas.

24. The video image annotation applying apparatus according to claim 2, wherein a plurality of items of visual information of the object as viewed, in part or wholly, in one direction or a plurality of directions are held as the visual feature information of the object stored in said storage means.

25. A vehicle guidance system adapted to be mounted on a vehicle for displaying a position of its own on a map displayed by a display device based on a GPS, comprising the video image annotation applying apparatus according to claim 2.

26. A method of recognizing a video image object, comprising the steps of:

inputting video image data and image capturing information which is information for determining an area where an image will be captured;

5 storing positional information which is information representing the position of an object and visual feature information which is information representing a numerical value of a visual feature of the object, in association with each other;

estimating the area where an image will be captured based on

10 the image capturing information;

matching the area where an image will be captured to a position represented by the positional information of the object which is stored;

extracting partial video image data which is either video image data of a partial area of the video image based on the video image data or is video image data of the entire video image, from the input video image;

generating visual feature information of the partial video image data;

comparing the visual feature information of the partial video image data and the stored visual feature information of the object to calculate a similarity therebetween; and

determining whether an image of an object is captured or not, based on the result of the matching and the calculated similarity.

27. A method of applying an video image annotation, comprising the steps of:

inputting video image data and image capturing information which is information for determining an area where an image will be captured;

storing positional information which is information representing the position of an object, visual feature information which is information representing a numerical value of a visual feature of the object, and additional information which is information added to the object, in association with each other;

estimating the area where an image will be captured based on the image capturing information;

matching the area where an image will be captured to a position represented by the positional information of the object which is stored;

extracting partial video image data which is either video image data of a partial area of the video image based on the video image data or is video image data of the entire video image, from the input video image;

generating visual feature information of the partial video image data;

comparing the visual feature information of the partial video image data and the stored visual feature information of the object to calculate a similarity therebetween; and

identifying an object which is contained in the video image, based on the result of the matching and the calculated similarity, and associating the identified object and the stored additional information with each other.

28. A video image object recognizing program adapted to be installed in a video image object recognizing apparatus for determining whether an object which is stored is contained as a subject in video image data or not, said video image object recognizing program to enable a computer to perform a process comprising the steps of:

storing, in a storage device, positional information which is information representing the position of an object and visual feature information which is information representing a numerical value of a visual feature of the object, in association with each other;

estimating an area where an image will be captured based on image capturing information which is information for determining the area where an image will be captured;

matching the area where an image will be captured to a position represented by the positional information of the object which is stored in said storage device;

extracting partial video image data which is either video image data of a partial area of the video image based on the video image data or is video image data of the entire video image, from input video image;

generating visual feature information of the partial video image

20   data;

comparing the visual feature information of the partial video image data and the stored visual feature information of the object to calculate a similarity therebetween; and

determining whether an image of an object is captured or not,

25   based on the result of matching and calculated similarity.

29. A video image annotation applying program adapted to be installed in a video image annotation applying apparatus for associating an object and information of an object which is stored with each other, said video image annotation applying program enabling a computer to perform a process

5   comprising the steps of:

storing, in a storage device, positional information which is information representing the position of an object, visual feature information which is information representing a numerical value of a visual feature of the object, and additional information which is information added to the object, in

10   association with each other;

estimating an area where an image will be captured based on image capturing information which is information for determining the area where an image will be captured;

matching the area where an image will be captured to a position

15   represented by the positional information of the object which is stored in said storage device;

extracting partial video image data which is either video image data of a partial area of the video image based on the video image data or is video image data of the entire video image, from input video image;

20   generating visual feature information of the partial video image data;

comparing the visual feature information of the partial video image data and the visual feature information of the object which is stored with each other to calculate a similarity therebetween; and

25   identifying an object which is contained in the video image, based on the result of matching and calculated similarity, and associating the identified object and the additional information which is stored with each other.